



## OPEN ACCESS

## EDITED BY

Prashant Vikram,  
Shree Guru Gobind Singh Tricentenary  
University, India

## REVIEWED BY

Sergio Alan Cervantes-Pérez,  
University of Arizona, United States  
Vignesh Muthusamy,  
Indian Agricultural Research Institute (ICAR),  
India  
Anupam Singh,  
Mahindra and Mahindra Limited, India

## \*CORRESPONDENCE

Manje Gowda,  
✉ m.gowda@cgiar.org

†These authors have contributed equally to  
this work

RECEIVED 10 December 2023

ACCEPTED 07 February 2024

PUBLISHED 21 February 2024

## CITATION

Ndlovu N, Kachapur RM, Beyene Y, Das B,  
Ogugo V, Makumbi D, Spillane C, McKeown PC,  
Prasanna BM and Gowda M (2024), Linkage  
mapping and genomic prediction of grain  
quality traits in tropical maize (*Zea mays* L.).  
*Front. Genet.* 15:1353289.  
doi: 10.3389/fgene.2024.1353289

## COPYRIGHT

© 2024 Ndlovu, Kachapur, Beyene, Das, Ogugo,  
Makumbi, Spillane, McKeown, Prasanna and  
Gowda. This is an open-access article  
distributed under the terms of the [Creative  
Commons Attribution License \(CC BY\)](#). The use,  
distribution or reproduction in other forums is  
permitted, provided the original author(s) and  
the copyright owner(s) are credited and that the  
original publication in this journal is cited, in  
accordance with accepted academic practice.  
No use, distribution or reproduction is  
permitted which does not comply with these  
terms.

# Linkage mapping and genomic prediction of grain quality traits in tropical maize (*Zea mays* L.)

Noel Ndlovu <sup>1,2†</sup>, Rajashekar M. Kachapur <sup>2,3†</sup>,  
Yoseph Beyene <sup>2</sup>, Biswanath Das <sup>2</sup>, Veronica Ogugo <sup>2</sup>,  
Dan Makumbi <sup>2</sup>, Charles Spillane <sup>1</sup>, Peter C. McKeown <sup>1</sup>,  
Boddupalli M. Prasanna <sup>2</sup> and Manje Gowda <sup>2\*</sup>

<sup>1</sup>Agriculture and Bioeconomy Research Centre, Ryan Institute, University of Galway, Galway, Ireland,  
<sup>2</sup>International Maize and Wheat Improvement Center (CIMMYT), Nairobi, Kenya, <sup>3</sup>University of  
Agricultural Sciences, Dharwad, Karnataka, India

The suboptimal productivity of maize systems in sub-Saharan Africa (SSA) is a pressing issue, with far-reaching implications for food security, nutrition, and livelihood sustainability within the affected smallholder farming communities. Dissecting the genetic basis of grain protein, starch and oil content can increase our understanding of the governing genetic systems, improve the efficacy of future breeding schemes and optimize the end-use quality of tropical maize. Here, four bi-parental maize populations were evaluated in field trials in Kenya and genotyped with mid-density single nucleotide polymorphism (SNP) markers. Genotypic (G), environmental (E) and G×E variations were found to be significant for all grain quality traits. Broad sense heritabilities exhibited substantial variation (0.18–0.68). Linkage mapping identified multiple quantitative trait loci (QTLs) for the studied grain quality traits: 13, 7, 33, 8 and 2 QTLs for oil content, protein content, starch content, grain texture and kernel weight, respectively. The co-localization of QTLs identified in our research suggests the presence of shared genetic factors or pleiotropic effects, implying that specific genomic regions influence the expression of multiple grain quality traits simultaneously. Genomic prediction accuracies were moderate to high for the studied traits. Our findings highlight the polygenic nature of grain quality traits and demonstrate the potential of genomic selection to enhance genetic gains in maize breeding. Furthermore, the identified genomic regions and single nucleotide polymorphism markers can serve as the groundwork for investigating candidate genes that regulate grain quality traits in tropical maize. This, in turn, can facilitate the implementation of marker-assisted selection (MAS) in breeding programs focused on improving grain nutrient levels.

## KEYWORDS

grain quality traits, quantitative trait loci, genomic prediction, tropical maize, sub-Saharan Africa

## 1 Introduction

Maize (*Zea mays* L.) ranks among the most prominent coarse cereal crops on a global scale, alongside rice and wheat (Erenstein et al., 2021). It is a widely cultivated crop, spanning over 170 countries and covering a vast area of 197 million hectares (FAOstat, 2022). In 2020, maize cultivation spanned approximately 43 million ha in Sub-Saharan

Africa (SSA), contributing to a substantial production of around 90 million metric tonnes. On a continental scale, Africa accounts for approximately 20.9% of the total global maize cultivation area. However, its contribution to global maize production is comparatively lower (~7.4%) (Prasanna et al., 2021). Indeed, maize yield growth varies across regions, with Africa lagging at 1.3%, while the United States of America leads at 2.0%, followed by Asia at 1.8% (Erenstein et al., 2022). Despite the widely reported low productivity per unit area, maize is the major source of food for more than 80% of the population in SSA (Prasanna et al., 2021) and meets more than 30% of their calorie requirement (Goredema-Matongera et al., 2021). Maize cuisine in this region is remarkably diverse, encompassing six distinct categories: whole-maize foods, wet-ground foods, snacks, bread, maize sourdough, and dumplings (Ekpa et al., 2018). Within this spectrum, stiff porridge, prepared by slowly adding maize flour to boiling water until it reaches the desired thickness - called *ugali* in Kenya and Tanzania, *nshima* in Zambia, *sadza* in Zimbabwe, and *mealiepap* in South Africa - stands out as a widely favoured culinary choice (De Groote and Kimenju, 2012). This staple, akin to rice in Asian cuisines, holds significant prominence and is widely consumed.

The substantial dependence of the SSA population on maize-based foods comes with inherent drawbacks, given that a significant proportion of the accessible maize varieties lack adequate levels of essential minerals. Ranum et al. (2014) indicated that while tropical maize is characterized by a high carbohydrate content (~72%), its grain protein content is modest, ranging from 9% to 10% and its fat content is around 4%. The region's reliance on maize for sustenance, particularly among low-income farming communities facing malnutrition, amplifies the urgency of addressing maize yield stagnation. Maize production in these communities is globally the lowest, standing at 2.1 t/ha. This can be attributed to various factors, with a considerable portion linked to the restricted availability of improved hybrid seeds (Quarshie et al., 2021). Other contributing factors include low-input agricultural practices, such as the application of only 8.0 kg of fertilizers per ha compared to the global average of 137 kg/ha (Das et al., 2019), small land holdings, limited access to mechanisation, sub-optimal post-harvest management, and challenges related to biotic and abiotic stresses. Some of these challenges impacting maize production in SSA have a negative effect on grain quality composition, notably the critical issue of low soil nitrogen stress. The application of nitrogenous fertilizer in smallholder farming systems within this region is severely constrained, reported by FAOSTAT (2023) at 4%. The effect of soil nitrogen on maize grain yield, composition and quality has been extensively investigated by numerous researchers (Worku et al., 2012; Biswas and Ma, 2016; Zhang et al., 2020; Ertiro et al., 2022; Ndlovu et al., 2022; Hammad et al., 2023). Although findings have varied, there is a consensus on the existence of genotypic differences in grain yield and composition among tropical maize genotypes grown under diverse management conditions.

The integration of genomic tools with field-based breeding techniques holds promise for enhancing nutrient composition in maize grain. Among these genomic techniques, quantitative trait loci (QTL) mapping stands out as a classical method for identifying genomic regions associated with traits of interest (Yang et al., 2020; Goering et al., 2021; Ren et al., 2022), even in the absence of prior genetic knowledge about the specific trait(s). Several studies have identified major QTLs associated with starch, protein, and oil

contents in maize grain. In 275 recombinant inbred lines (RILs), Lu et al. (2022) identified 11 QTLs affecting kernel protein content some linked to the *Zm00001d002625* gene which encodes an S-adenosyl-L-methionine-dependent methyltransferase superfamily protein. Guo et al. (2013) identified nine unconditional QTLs across all chromosomes (excluding chr 3 and 7) for oil content and eight unconditional QTLs distributed across all chromosomes except chromosomes 4 and 8 for starch content. Liu et al. (2008) identified a total of eighteen QTLs for grain quality traits across diverse soil nitrogen regimes and locations. Among these, seven QTLs were associated with oil content, six with protein content, and five with starch content. Ndlovu et al. (2022) also identified multiple QTLs for maize grain quality traits using multiple biparental populations under different soil nitrogen levels. Application of genomic selection for different traits in maize revealed moderate to high prediction accuracies (Gowda et al., 2018; 2021; Ertiro et al., 2020; Beyene et al., 2021). Ndlovu et al. (2022) applied genomic selection on maize grain quality traits under optimum and low soil N management and observed moderate to high accuracies in both diversity panel and biparental populations. In the preceding studies, empirical evidence has substantiated that the integration of traditional breeding methodologies with linkage mapping, coupled with genomic selection, markedly enhances the efficiency of improving nutritional quality traits in maize.

Though many studies reported several genomic regions associated with grain quality traits, additional sources of genetic variation exist in unexplored maize populations. On the flip side, the rapid acquisition of high-quality phenotypic data to facilitate genomic analyses for grain-quality traits continues to be a resource-demanding endeavour. Indeed, the conventional phenotyping methods for assessing grain quality traits are characterized by their labour-intensive nature and time-consuming processes. On a positive note, the expanded utilization of Near Infrared Reflectance Spectroscopy (NIRS) has enhanced the capabilities of nutritional profiling studies for maize grain. NIRS is a fast, reliable, and non-destructive method that is being employed in plant phenotyping assessment of maize kernel starch, protein and oil content (Ndlovu et al., 2022). In this study, we used NIRS to measure the main nutritional quality traits in four biparental maize populations and used mid-density single nucleotide polymorphism (SNP) markers for both linkage mapping and genomic prediction. Our study was designed to accomplish the following objectives: (i) to assess the genetic variation in grain quality traits (i.e., protein, oil and starch) among tropical maize populations; (ii) to identify significant QTLs associated with grain quality traits in tropical maize populations tested across multiple environments, and (iii) to assess the potential of utilizing genomic selection for the improvement of grain quality traits in tropical maize.

## 2 Materials and methods

### 2.1 Plant materials

One doubled haploid (DH) and three F<sub>3</sub> tropical maize populations comprising 110, 271, 333 and 158 lines, respectively,

TABLE 1 Details of the maize populations used in this study.

Population	Pedigree	Population size
DH Population 1	CML494×CML550	110
F <sub>3</sub> Population 2	CKL05017×CML536	276
F <sub>3</sub> Population 3	CML494×CML550	315
F <sub>3</sub> Population 4	VL081452×VL058589	158

were developed from four bi-parental crosses (Table 1). The populations used in this study were also used for mapping agronomic traits in our earlier studies (Ertiro et al., 2020; Ndlovu et al., 2022; Kimutai et al., 2023). All the parental lines we used in this study exhibited variations in grain quality traits (oil, protein and starch content) and are adapted to mid-altitude regions (1,000–1,500 m above sea level (MASL)) of SSA. Studied populations were test-crossed with an appropriate tester from the opposite heterotic group for phenotypic evaluations. The performance of parental lines and selected commercial hybrids (Supplementary Table S1) and the progenies for each population are listed (Supplementary Table S2) for all quality traits. Each population along with its parents were planted at the Kiboko Maize Research Station, Kiboko, Kenya. Kiboko Maize Research Station is situated between 37.7235°E longitude and 2.2172°S latitude, at an elevation of 975 MASL. The annual temperature ranges from 16.0°C to 33.6°C, and the rainfall varies from 545 to 629 mm annually across the two studied seasons. This location lies in a hot, semi-arid region and the soils are well-drained, dark reddish brown to dark red, friable sandy clay to clay (Acric-Rhodic Ferrosols) developed from undifferentiated basement system rocks, predominantly banded gneisses (Ertiro et al., 2022). The DH pop 1 and F<sub>3</sub> pop 2 were evaluated at Kiboko in the main rainy season for two seasons, while F<sub>3</sub> pop 3 and F<sub>3</sub> pop 4 were evaluated for one season.

## 2.2 Field trial and data collection

All four biparental populations were evaluated using an alpha lattice incomplete block design with two replications. Single row plots, measuring 5 m long at a row spacing of 0.75m, were sown. Trial plots were top-dressed with urea fertilizer at the rate of 138 kg N per hectare 3 weeks post-planting. All trials were irrigated as required to avoid any moisture stress. Trials were kept weed-free and other established standard agronomic practices were followed. Data pertaining to the target traits were systematically collected by selecting ten plants from the midsection of the plot rows. This included plant height (PH, centimetres), anthesis date (AD, days), anthesis silking interval (ASI, days), ear height (EH, centimetres) and ear position (EPO, ratio of EH/PH). For grain yield (GY) assessment, the shelled grains were quantified in kilograms (kg) and subsequently converted to tons per hectare, considering a moisture content of 12.5% (GY in t/ha).

Grain texture was measured on a 1 to 5 scale (where 1 = flint, 2 = semi-flint, 3 = intermediate, 4 = semi-dent and 5 = dent). Kernel weight (Kwt) was determined by measuring 100 randomly

selected seeds per line/family per replication, with measurements recorded in grams. Following harvest, seeds for each genotype were separated to facilitate grain nutrient analyses. Grain quality traits, specifically oil, protein, and starch contents, were quantified using a FOSS Infratec TM 1241. The analysis involved 500 g samples of maize grain obtained from each plot, and the results were reported as a percentage of whole grain. Five 100-g subsamples were assayed and the mean reading for each parameter was reported per plot. The reflectance spectra were collected in a range of 400–2,500 nm with 10 nm intervals in the near-infrared reflectance (NIR) region.

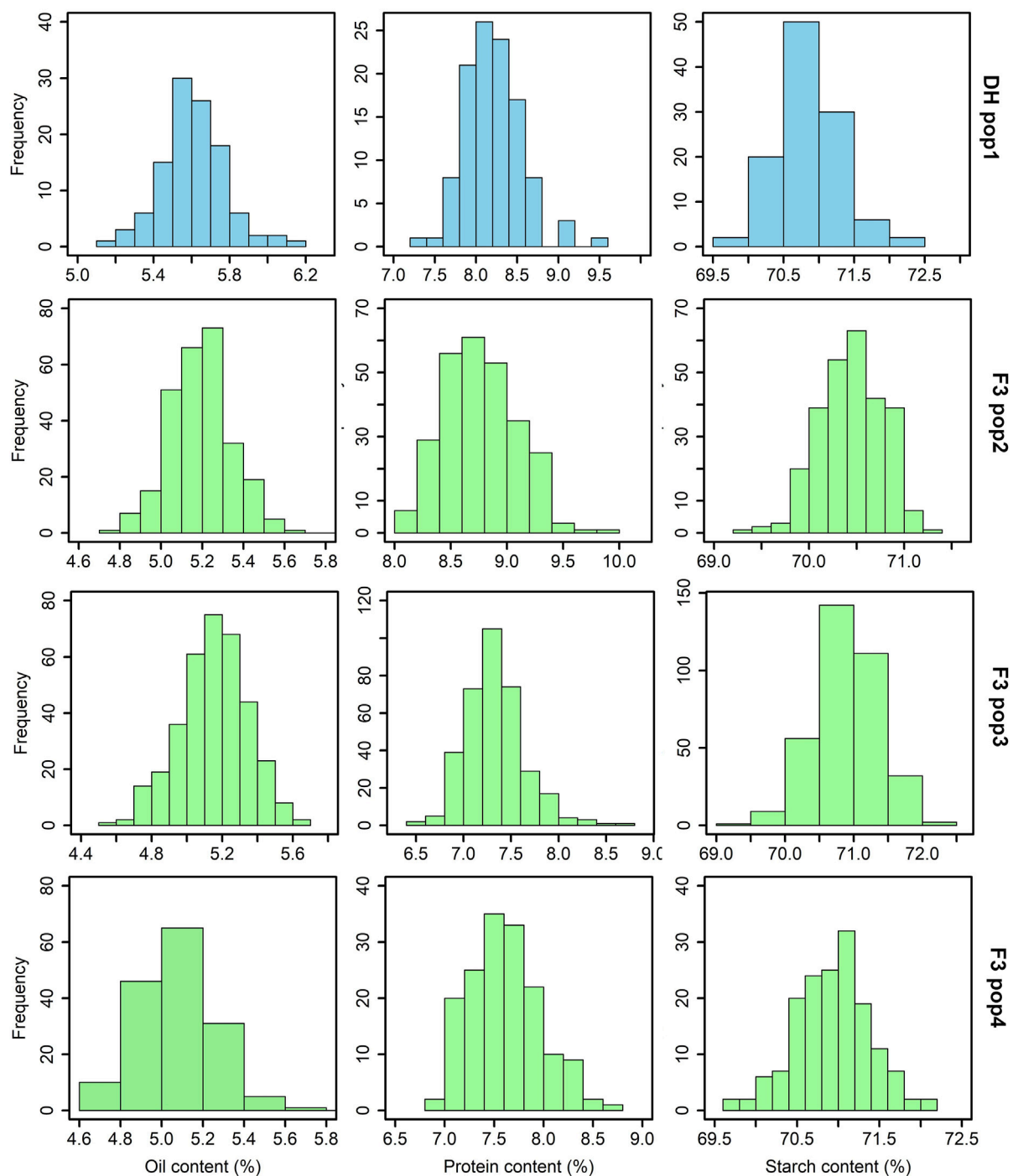
Analyses of variance (ANOVA) for grain traits, including oil content, protein content, starch content, kernel weight, and grain texture, within each biparental population across different seasons were performed using the META-R (Alvarado et al., 2020) and ASREML-R (Gilmour et al., 2002). Best Linear Unbiased Estimators (BLUEs) were computed utilizing a mixed model, wherein genotype entries were treated as fixed effects, while the remaining terms were treated as random. In the estimation of broad-sense heritability, all terms were considered as random effects. Broad sense heritability was estimated by the formula:

$$h^2 = \sigma_G^2 / (\sigma_G^2 + \sigma_{GE}^2/E + \sigma_e^2/Er)$$

Where  $\sigma_G^2$  is the genotypic variance,  $\sigma_{GE}^2$  is the genotypic by environment interaction (GEI),  $\sigma_e^2$  is the error variance, E is the number of environments or seasons, and r is the number of replications in each trial. The phenotypic and genotypic correlations among traits were evaluated as described by Hirel et al. (2007).

## 2.3 Genotypic analysis

DNA was extracted from bulked young leaves of the studied tropical maize lines following the CTAB Method (CIMMYT, 2005). Genotyping of the F<sub>3</sub> populations was performed using the Illumina Maize SNP1500 Bead Chip, which utilizes evenly spaced SNPs to comprehensively cover the maize genome (Ganal et al., 2011). The above task was performed at the LGC genomic labs in the United Kingdom (<https://www.lgcgroup.com/genotyping/>). In the case of the tested DH population, lines were genotyped using the Genotyping-by-Sequencing (GBS) approach. The obtained data underwent filtration using TASSEL software, with criteria set at > 0.10 Minor Allele Frequency (MAF), <5% heterozygosity, and a minimum count of 90% of the total size (Bradbury et al., 2007; Sironik et al., 2019). In all the populations, homozygous marker loci for both parents and uniformly distributed polymorphic markers between parents were retained. Linkage maps in all four populations were constructed using QTL IciMapping version 4.1 software (Meng et al., 2015). After removing the distorted markers, finally, we used 1,007, 452, 231, and 387 high-quality SNPs in DH pop 1, F<sub>3</sub> pop 2, F<sub>3</sub> pop 3 and F<sub>3</sub> pop 4, respectively. In brief, the linkage map was constructed by using these SNPs, and by selecting the most significant markers using stepwise regression. A likelihood ratio test was used to calculate the logarithm of odds (LOD) for each marker at a score of >3 with a 30 cM maximum distance between two loci. The recombination frequency between linked loci was



**FIGURE 1**  
Phenotypic distribution for grain quality traits evaluated across DH and F<sub>3</sub> tropical maize populations. The sky blue and light green colour plots represent the grain quality trait measurements for DH and F<sub>3</sub> populations, respectively. DH pop 1 = CML494×CML550; F<sub>3</sub> pop 2 = CKL05017×CML536; F<sub>3</sub> pop 3 = CML494×CML550; and F<sub>3</sub> pop 4 = VL081452×VL05858.

transformed into cM (CentiMorgan's units) using Kosambi's mapping function (Kosambi, 1944).

For each population, best linear unbiased predictors (BLUPs) across seasons were used to detect QTLs based on Inclusive interval mapping (ICIM). The quantification of phenotypic variation attributed to individual QTLs and the cumulative variation explained by the

aggregate presence of all QTLs was conducted. QTL names were constructed by starting with the letter 'q' to indicate QTL, followed by an abbreviation of the trait name, the corresponding chromosome number, and the marker position (Ribaut et al., 1997).

Genomic prediction (GP) analysis was performed in R program version 4.2.1 (R Core Team 2023). GP was applied to

TABLE 2 Estimates of means, components of genotypic ( $\sigma^2_G$ ), genotype  $\times$  environment interaction ( $\sigma^2_{G \times E}$ ), error variances ( $\sigma^2_e$ ) and heritability ( $h^2$ ) for four biparental populations evaluated at Kiboko in Kenya for grain quality traits (protein content, starch content, oil content and grain texture).

Populations	Grain quality traits			
	Oil content	Protein content	Starch content	Texture/Kernel weight <sup>1</sup>
<b>DH pop 1 - CML494<math>\times</math>CML550</b>				
$\sigma^2_G$	0.02*	0.03*	0.04**	4.37*
$\sigma^2_{G \times E}$	0.01	0.03*	0.02*	-
$\sigma^2_e$	0.03	0.24	0.41	21.30
$h^2$	0.68	0.29	0.26	0.31
LSD <sub>5%</sub>	0.22	0.38	0.49	5.07
CV (%)	3.23	5.94	1.01	7.06
<b>F<sub>3</sub> pop 2 - CKL5017<math>\times</math>CML536</b>				
$\sigma^2_G$	0.02*	0.04**	0.04**	0.07**
$\sigma^2_{G \times E}$	0.01	0.06**	0.01*	0.04*
$\sigma^2_e$	0.04	0.11	0.21	0.19
$h^2$	0.58	0.42	0.42	0.67
LSD <sub>5%</sub>	0.22	0.44	0.46	0.43
CV (%)	3.66	3.85	0.69	19.76
<b>F<sub>3</sub> pop 3 - CML494<math>\times</math>CML550</b>				
$\sigma^2_G$	0.02*	0.04*	0.07**	3.47**
$\sigma^2_e$	0.03	0.09	0.27	14.84
$h^2$	0.57	0.47	0.34	0.35
LSD <sub>5%</sub>	0.26	0.42	0.60	3.11
CV (%)	3.55	4.18	0.73	7.33
<b>F<sub>3</sub> pop 4 - VL081452<math>\times</math>VL058589</b>				
$\sigma^2_G$	0.02*	0.04*	0.04*	11.50**
$\sigma^2_e$	0.04	0.15	0.22	30.42
$h^2$	0.50	0.35	0.27	0.43
LSD <sub>5%</sub>	0.30	0.46	0.51	7.38
CV (%)	4.16	5.17	0.80	7.89

CV- coefficient of variation, LSD-least significant difference, \*, \*\* significant at  $p=0.05$  and  $0.01$  level, respectively.<sup>1</sup> Grain texture was recorded in F<sub>3</sub> pop 2 whereas kernel weight was recorded in the other three populations.

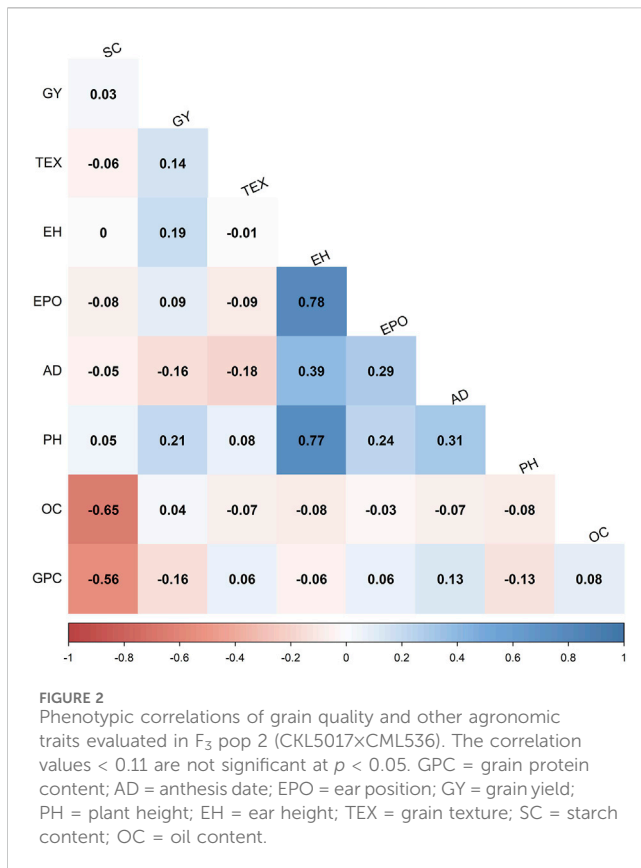
each F<sub>3</sub> and DH population to find out the prediction accuracy of grain quality traits. This was done using the RR-BLUP model (Zhao et al., 2012; Crossa et al., 2017). BLUEs across seasons for each of the studied populations were used for the analysis. In GP analysis, polymorphic SNPs between the parents of each population, comprising 1,007, 452, 231, and 387 SNPs in DH pop 1, F<sub>3</sub> pop 2, F<sub>3</sub> pop 3, and F<sub>3</sub> pop 4, respectively, were used. A five-fold cross-validation approach was employed, specifically utilizing a 'within population' strategy where both training and estimation sets originated from within each bi-parental population. For each trait in each population, 100 iterations were performed to divide the data into training and estimation sets.

## 3 Results

### 3.1 Phenotypic analyses

The analysis of grain quality traits revealed significant variability within the studied tropical maize populations for each trait. The distribution of these traits closely followed the expected normal distribution pattern, as demonstrated by the histogram plots (Figure 1). Across the DH and F<sub>3</sub> populations, the mean protein content varied from 6.8%–10%, oil content from 4.5%–6.2% and starch content ranged from 69.5%–72.5%. Among these populations, F<sub>3</sub> pop 2 had the highest protein content (10%) followed by DH pop 1 (9.6%). F<sub>3</sub> pop 3 and F<sub>3</sub> pop 4 recorded





grain protein levels <9%. For oil content, DH pop 1 (6.2%) exhibited the highest values, followed by F<sub>3</sub> pop 4 (5.8%). For starch content, the highest levels were recorded in DH pop 1 and F<sub>3</sub> pop 3 (both with 72.5%).

Genotypic variance ( $\sigma^2_G$ ) was significant at  $p \leq 0.05$  for grain quality traits (oil, protein, starch content, kernel weight and grain texture) across all the studied genotypes (Table 2). Genotype-environment interaction effects ( $\sigma^2_{G \times E}$ ) were significant for protein and starch content in DH pop 1 and F<sub>3</sub> pop 2. In general, kernel weight consistently exhibited elevated coefficients of variation (CVs) within both DH and F<sub>3</sub> populations. When comparing the same trait across the studied populations, the CVs for the F<sub>3</sub> pop 4 were marginally higher than those for other populations. Nonetheless, substantial variations across the genotypes and traits were still evident.

Across the segregating populations, the highest and lowest values of broad sense heritabilities were observed for oil content (0.68) and starch content (0.26) in DH pop 1. Overall broad-sense heritabilities ranged from low to high for all grain quality traits studied—with oil content (0.68) and grain texture (0.67) having the highest values. Generally, oil content (0.50–0.68) had moderate to high broad-sense heritabilities across the studied populations. Whereas protein (0.29–0.47) and starch (0.26–0.42) contents, on the other hand, recorded low to moderate values of heritability across the populations. Likewise, for kernel weight trait the heritability values ranged from low (0.31 in DH pop 1) to high (0.43 in F<sub>3</sub> pop 4).

To understand the interrelations among grain quality traits, Pearson's correlations of BLUP values were computed within the F<sub>3</sub>

pop 2 (Figure 2). Grain protein content was negatively correlated with starch content ( $r = -0.56^{**}$ ), grain yield ( $r = -0.16^{**}$ ) and plant height ( $r = -0.13^*$ ). In the same population, grain yield had a weak but significant positive correlation with grain texture ( $r = 0.14^*$ ), plant height ( $r = 0.21^{**}$ ) and ear height ( $r = 0.19^{**}$ ). Grain oil content was negatively correlated with starch content ( $r = -0.65^{**}$ ). Grain texture and anthesis date were also negatively correlated ( $r = -0.18^{**}$ ). The documented correlations between grain quality and associated traits in diverse maize populations present valuable insights for making informed decisions in genotypic selection. These findings suggest a complex trait architecture, wherein grain quality traits display significant interactions (+/–) with one another.

### 3.2 QTLs associated with grain quality traits

The maize populations evaluated in this study were also tested in earlier studies (Gowda et al., 2018; Kimutai et al., 2023), where the linkage maps information is included. In brief, the map length of each population was 2,970, 1,650.28, 906.83, and 2,169.97 cM from 1,007, 452, 202, and 387 polymorphic SNPs for DH pop 1 and F<sub>3</sub> pop 2, 3, and 4, respectively. Across the ten maize chromosomes, a total of 63 significant QTLs were identified for oil content (13), grain protein content (7), starch content (33), grain texture (8) and kernel weight (2) (Table 3). The identified QTLs were distributed across 10 chromosomes. Table 3 contains comprehensive information about the identification and locational specifics of the discerned QTLs, as well as their respective genetic effects. QTLs for oil content were found in all chromosomes except chromosomes 2, 7 and 10. For protein content, QTLs were only discovered on chromosomes 1, 2, 3, 5 and 6. QTLs associated with starch content were found in all chromosomes except on chromosomes 9 and 10. For grain texture, only chromosomes 6 and 10 had no QTLs. For kernel weight, two QTLs on chromosomes 4 and 10 were found in F<sub>3</sub> pop3.

In DH pop1, QTL analysis revealed a total of 6 QTLs for the studied grain quality traits on chromosomes 3, 6, 8, 9 and 10 (Table 3). The highest number of QTLs were discovered for starch content in F<sub>3</sub> pop 2 ( $n = 22$ ). For the same population, 8 and 4 QTLs discovered were associated with grain texture and protein content respectively. No QTLs were detected for protein content in F<sub>3</sub> pop 4. A comparison of the QTLs across DH and F<sub>3</sub> populations revealed that several QTLs overlapped for some of the grain quality traits. In Chromosome 6, the region between 89.81 and 112.9 Mb had QTLs for both oil and starch content. In chromosome 2, QTLs associated with protein content and starch content were detected in the region between 188.45 and 233.61 Mb (Table 3).

The phenotypic variance explained (PVE) for all detected QTLs associated with grain quality traits ranged from 0.81% to 26.85% (Table 3). The most extensive range of PVE was observed in F<sub>3</sub> pop 2 (0.81%–22.46%) for starch content. On the other hand, a narrow PVE range was observed for oil content (2.49%–3.51%) in F<sub>3</sub> pop 3. Significant QTLs that explained more than 10% of the PVE (major effects), was identified for oil content ( $qOC3-146$  (18.06%),  $qOC6-102$  (26.85%),  $qOC2-185$  (25.93%) and  $qOC2-185$  (25.93%)), protein content ( $qGPC6-30$  (19.45%)), starch content ( $qSC1-20$  (22.46%),  $qSC4-10$  (10.55%),  $qSC5-10$  (12.99%) and  $qSC5-55$  (19.75%)) and grain texture ( $qGT3-9$  (14.59%) and  $qGT5-88$  (11.29%)). In DH pop1, additive effects are high for two major effect QTL and the

TABLE 3 Analysis of markers associated with grain quality traits, allele substitution (a) effects, and the total phenotypic variance of the joint linkage association mapping based on combined DH and F3 maize populations.

QTL name	Chr	Position (cM)	LOD	PVE (%)	Add	Dom	Total PVE (%)	Flanking markers		Physical position (Mbp)
DH pop 1 - CML494×CML550										
Oil content										
<i>qOC3-146</i>	3	150	6.3	18.06	0.05	-	30.72	S3_145603187	S3_149385251	145.60–149.38
<i>qOC6-102</i>	6	50	8.8	26.85	0.06	-		S6_101991919	S6_103359585	101.99–103.35
<i>qOC8-16</i>	8	197	2.98	8.79	-0.03	-		S8_15877080	S8_19122987	15.87–19.12
<i>qOC9-10</i>	9	48	3.47	9.58	0.03	-		S9_8416672	S9_10276690	8.41–10.27
Grain protein content										
<i>qGPC9-144</i>	9	162	2.95	11.88	-0.02	-	10.75	S9_143801329	S9_149382254	143.80–149.38
Starch content										
<i>qSC10-20</i>	10	68	2.81	7.44	-0.02	-	9.75	S10_11602996	S10_38530632	11.60–38.53
F3 pop 2 - CKL5017×CML536										
Oil content										
<i>qOC1-64</i>	1	165	3.84	7.67	-0.06	0.04	15.53	PZA00455.14	PHM 1968.22	63.80–168.71
<i>qOC6-74</i>	6	59	4.59	7.13	-0.06	0.02		PZB01009.1	PZA00942.2	72.95–89.81
<i>qOC6-90</i>	6	70	3.74	5.6	0.05	0		PZA00942.2	lac1.3	89.81–112.90
Protein content										
<i>qGPC1-105</i>	1	159	5.94	9.14	-0.14	0	17.61	PZA00939.1	PZA00455.14	104.96–168.71
<i>qGPC2-227</i>	2	46	2.74	6.19	-0.02	-0.16		PZD00022.5	PZA02266.3	226.34–233.61
<i>qGPC2-10</i>	2	207	3.42	4.89	0.09	0.04		PHM6111.5	PZA00613.22	3.50–21.42
<i>qGPC6-30</i>	6	28	3.05	19.45	-0.17	-0.04		PZD00072.2	PZA00440.1	21.59–71.99
Starch content										
<i>qSC1-20</i>	1	16	5.6	22.46	0.04	0.1	17.96	PHM11114.7	PZA01456.2	12.76–157.65
<i>qSC1-105</i>	1	158	3.84	0	0.01	0		PZA00939.1	PZA00455.14	104.96–168.71
<i>qSC1-21</i>	1	303	3.59	8.81	-0.03	-0.3		PHM574.14	PZA02094.9	15.33–59.27
<i>qSC2-189</i>	2	1	2.96	2.82	0.02	-0.01		PZA01552.1	PZD00022.5	188.45–233.61
<i>qSC2-194</i>	2	77	3.11	2.98	0.03	0.01		PZA02964.7	PHM14412.4	193.67–201.69
<i>qSC2-189</i>	2	79	3.2	2.2	0.03	0.02		PHM7953.11	PZA00803.3	189.50–189.60
<i>qSC3-171</i>	3	20	2.55	4.5	0.04	0.03		PZA03735.1	PHM17210.5	170–91 - 173.26
<i>qSC3-121</i>	3	38	3.28	0.1	0.03	0.01		PHM1745.16	PZA00363.7	120.53–129.09
<i>qSC3-40</i>	3	44	9.03	4.47	0.04	0		PZA00707.9	PZA00380.10	38.20–98.45
<i>qSC3-38</i>	3	47	2.67	0.85	-0.03	0.01		PHM2343.25	PZA00297.2	26.85–38.83
<i>qSC4-10</i>	4	113	2.61	10.55	0.01	0.07		PZA02289.2	PZA00436.7	6.55–177.80
<i>qSC4-20</i>	4	143	3.92	6.17	-0.04	0.01		PHM259.11	PZA00704.1	16.30–114.90
<i>qSC5-10</i>	5	11	3.5	12.99	0.03	-0.2		PHM5484.22	PHM3137.17	8.18–20.83
<i>qSC5-2</i>	5	55	2.97	3.72	-0.03	0.02		PZA00963.3	PZA00818.1	1.0–87.11
<i>qSC5-119</i>	5	113	3.15	4	-0.03	-0.02		PZA02164.16	PZA01365.1	117.35–138.85
<i>qSC6-90</i>	6	60	5.14	2.56	0.03	-0.01		PZA00942.2	lac1.3	89.81–112.90

(Continued on following page)

TABLE 3 (Continued) Analysis of markers associated with grain quality traits, allele substitution (a) effects, and the total phenotypic variance of the joint linkage association mapping based on combined DH and F3 maize populations.

QTL name	Chr	Position (cM)	LOD	PVE (%)	Add	Dom	Total PVE (%)	Flanking markers		Physical position (Mbp)
<i>qSC6-113</i>	6	71	5.18	0.81	-0.03	0		lac1.3	PZA00571.1	112.90–113.26
<i>qSC6-125</i>	6	88	5.05	3.39	-0.04	0.05		PZB00414.2	PZA02328.5	124.71–130.59
<i>qSC7-15</i>	7	35	4.82	2.44	0.02	0		PZB00752.1	PZA02872.1	11.07–113.74
<i>qSC8-10</i>	8	9	9.85	1.24	0.01	0		PHM2487.6	PZA02955.3	8.40–15.16
<i>qSC8-105</i>	8	39	5.65	1.93	0	0.03		PZA02566.1	PHM934.19	104.22–109.38
<i>qSC8-123</i>	8	47	3.08	5.06	0.02	0.06		PZA01049.1	PHM4757.14	122.03–144.66
<b>Grain Texture</b>										
<i>qGT1-43</i>	1	157	4.5	4.76	0.06	-0.03	40.55	PZA00962.1	PZA00939.1	42.32–104.96
<i>qGT1-55</i>	1	304	2.53	5.52	-0.3	-0.53		PHM574.14	PZA02094.9	15.33–59.27
<i>qGT3-9</i>	3	68	11.5	14.59	0.11	-0.03		PZA00508.2	PZA01765.1	8.86–11.33
<i>qGT4-10</i>	4	123	7.21	8.03	-0.08	-0.02		PZA02289.2	PZA00436.7	6.55–177.80
<i>qGT5-88</i>	5	112	6.33	11.29	0.1	0.02		PZA01693.1	PZA02164.16	87.11–117.35
<i>qGT7-82</i>	7	91	2.73	7.71	-0.02	0.56		PZA01933.3	PHM3435.6	80.61–141.80
<i>qGT9-87</i>	9	22	2.76	3.49	0.06	0		PHM1766.1	PZA03235.1	86.54–107.26
<i>qGT9-80</i>	9	32	6.68	7.06	-0.08	0		PZA03235.1	PZA00225.8	76.31–86.54
<b>F3 pop 3 - CML494×CML550</b>										
<b>Oil content</b>										
<i>qOC3-06</i>	3	77	5.53	3.51	0.04	0	15.53	PZA01765.1	PHM12859.7	5.16–8.86
<i>qOC5-05</i>	5	78	3.47	2.49	0.03	0.01		PHM14671.9	PHM3096.19	3.83–5.50
<b>Grain protein content</b>										
<i>qGPC3-175</i>	3	30	2.89	4.21	0.1	-0.04	17.61	PZA00538.15	PHM17210.5	170.91–199.68
<i>qGPC5-10</i>	5	53	5.23	6.16	-0.12	0.03		PZA03340.2	PHM14671.9	5.50–19.33
<b>Starch content</b>										
<i>qSC1-227</i>	1	38	3.79	2.38	0.02	0.02	17.96	PHM5293.11	PZA00381.4	226.24–232.45
<i>qSC1-192</i>	1	67	2.56	3.57	0.04	0.03		PHM5480.17	PZA00425.11	191.75–209.98
<i>qSC3-175</i>	3	32	7.8	1.19	-0.03	0		PZA00538.15	PHM17210.5	170.91–199.68
<i>qSC3-06</i>	3	76	3.6	3.57	-0.04	0.02		PZA01765.1	PHM12859.7	5.16–8.86
<i>qSC5-70</i>	5	49	11.2	2.38	0.02	-0.01		PHM13675.18	PHM 1870.20	61.25–71.08
<i>qSC5-35</i>	5	51	13.8	1.19	0.02	-0.02		PZA01530.1	PZA00934.2	31.50–39.24
<i>qSC6-121</i>	6	65	3.01	1.19	-0.02	-0.01		PZA00473.5	PZA02673.1	117.06–135.75
<i>qSC8-10</i>	8	16	3.64	1.19	0.01	-0.03		PHM2487.6	PZA03178.1	8.41–11.88
<b>Kernel weight</b>										
<i>qKw4-170</i>	4	62	3.02	4.54	-0.31	0.10	9.27	PHM3155.14	PZA01187.1	168.25–175.15
<i>qKwt10-100</i>	10	1	3.84	4.50	0.33	0.05		PZB01358.1	PHM229.15	92.58–107.42
<b>F3 pop 4 - VL081452×VL058589</b>										
<b>Oil content</b>										
<i>qOC2-185</i>	2	207	11.6	25.93	-0.011	0.64	15.53	PZA02170.1	PHM5060.12	184.40–231.77

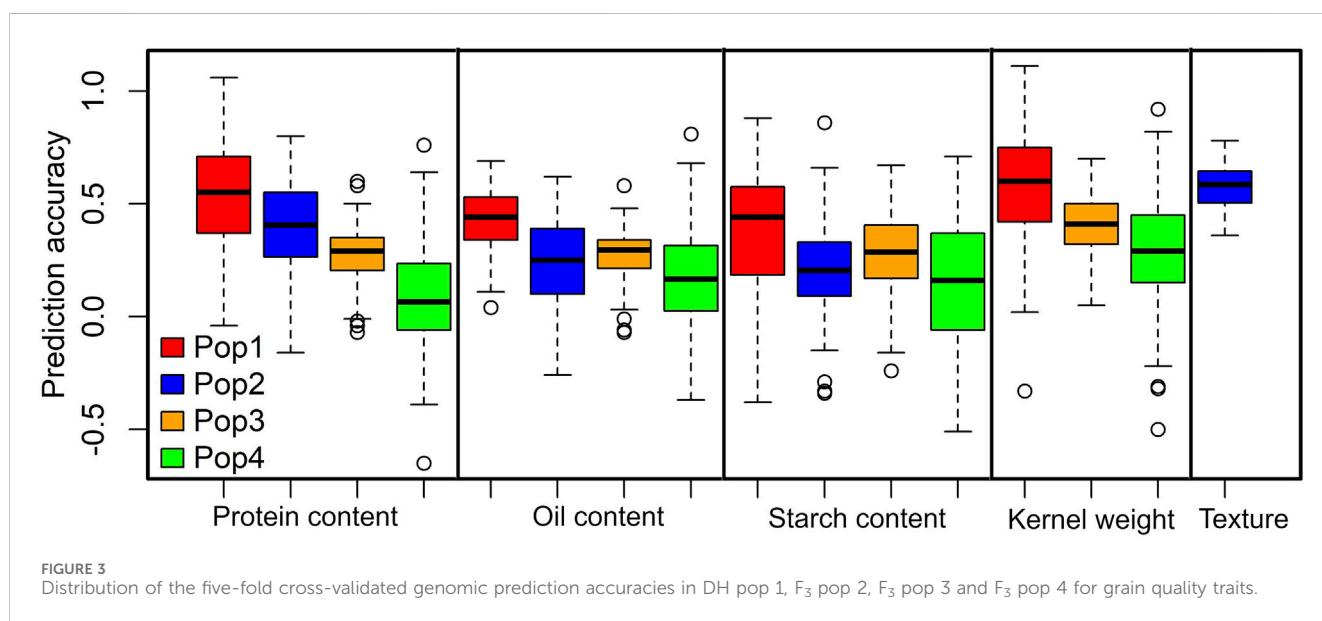
(Continued on following page)



TABLE 3 (Continued) Analysis of markers associated with grain quality traits, allele substitution (a) effects, and the total phenotypic variance of the joint linkage association mapping based on combined DH and F<sub>3</sub> maize populations.

QTL name	Chr	Position (cM)	LOD	PVE (%)	Add	Dom	Total PVE (%)	Flanking markers		Physical position (Mbp)
<i>qOC4-230</i>	4	89	3.26	5.98	0.03	0.001		PZA00521.3	PZA01367.2	226.98–243.38
<i>qOC4-170</i>	4	126	3.67	6.74	0.041	−0.01		PHM 2006.57	PZA01289.1	167.40–171.71
<i>qOC6-70</i>	6	126	3.78	6.97	0.037	−0.03		PZA00006.17	PHM14522.5	66.16–73.59
<b>Starch content</b>										
<i>qSC5-55</i>	5	196	3.65	19.74	0.095	−0.46	19.81	PZA02207.1	PHM2769.43	51.25–59.62
<i>qSC7-121</i>	7	43	2.5	7.34	−0.091	0.201		PHM9162.135	PZA02959.14	120.22–133.75

Chr—chromosome; PVE, phenotypic variance explained; add—additive effects; Dom—dominance effects.



favourable alleles are contributed from parent CML550. Whereas for the major effect QTL identified for grain protein content, the favourable alleles are contributed from parent CML494 (Table 3). In F<sub>3</sub> pop 2, the favourable alleles for major effect QTL for protein content was contributed from parent CKL5017, whereas for starch content and grain texture, parent CML536 contributed.

### 3.3 Prediction accuracies of grain quality traits in DH and F<sub>3</sub> tropical maize populations

To estimate the prediction accuracy for each of the studied grain quality traits, we used the RR-BLUP model (Figure 3). Overall, the prediction accuracies across the populations were moderate to high for the studied grain quality traits. The average prediction accuracies were higher for kernel weight in DH pop 1 (0.58) followed by grain texture (0.57). The prediction accuracies for protein content were 0.52, 0.41, 0.28 and 0.08 for DH pop 1, F<sub>3</sub> pop 2, F<sub>3</sub> pop 3 and F<sub>3</sub> pop 4, respectively (Figure 3). For oil content prediction accuracies were 0.43, 0.25, 0.28 and 0.17, whereas for starch content prediction

accuracies were 0.38, 0.21, 0.28 and 0.15 in DH pop 1, F<sub>3</sub> pop 2, F<sub>3</sub> pop 3, and F<sub>3</sub> pop 4, respectively. Interestingly, DH pop 1 outperformed F<sub>3</sub> maize populations in terms of overall trait genomic prediction accuracy for all traits.

## 4 Discussion

### 4.1 The positive and negative correlations among grain quality traits indicate their potential for simultaneous improvement through phenotypic selection

The significant variability observed within the four tropical maize populations for the studied grain quality traits (Figure 1; Table 2) underscores the genetic diversity present within the studied maize genotypes, emphasizing the potential for targeted breeding efforts to capitalize on this diversity and enhance grain quality traits in the subsequent generations. In an earlier study by Ndlovu et al. (2022), a pronounced range of variability in grain quality traits in maize was documented, despite their study being conducted under

conditions of low-N-induced stress. In our study, the tested biparental maize populations exhibited a typical grain quality trait composition with moderate protein (6.8%–10%) and oil (4.5%–6.2%) levels, and a substantial starch content reaching 72.5%. These values are consistent with the general nutritional profile of maize reported in earlier studies (Dei, 2017; Ray et al., 2019; Álvarez-Iglesias et al., 2021; Ertiro et al., 2022; Langyan et al., 2022; Wang et al., 2023). Targeted breeding could optimize the grain quality traits to meet the requirements of the food and feed industry.

The observed significant genotypic variance for grain quality traits underscores the substantial genetic diversity present within the tropical maize populations, presenting opportunities for trait enhancement through selective breeding. Furthermore, the significant G×E interactions observed for traits like protein and starch content underscore the importance of incorporating environmental factors into breeding strategies for these traits. These findings align with those of Katsenios et al. (2021) and Ndlovu et al. (2022), particularly regarding the grain protein content in maize. The results of these studies indicate that grain quality is influenced by a variety of environmental conditions. In this respect, to enhance grain quality traits, breeders and seed growers should consider cultivating maize lines in optimal environments. Despite the observed significant G×E interactions for these traits, environmental variance across the studied grain quality traits was not significant. This indicates that genetic factors predominantly govern the trait variations, minimizing the role of external environmental conditions.

Broad sense heritabilities for all the studied grain quality traits ranged from low to high, indicating the varying degrees of genetic influence on the trait variability. Grain oil content demonstrated moderate to high heritability, suggesting that a significant portion of the phenotypic variation was attributed to genetic factors, making it feasible for recurrent selection approaches. Conversely, grain protein and starch content exhibited a range of heritability from low to high, indicating a complex interplay of genetic and environmental factors. The significant negative correlations observed among all grain quality traits (Figure 2) indicate an inverse relationship, suggesting that an increase in one of the traits can lead to a decrease in the other trait. This result has implications for trait selection in breeding programs for improved grain quality in tropical maize, highlighting the need to carefully consider trade-offs and prioritize traits based on the set breeding objectives.

## 4.2 Grain quality traits in tropical maize are controlled by multiple QTLs

Grain quality traits in maize are characterized by their complexity as quantitative traits, governed by a combination of both major and minor genetic effects (Zheng et al., 2021). In the present study, we used linkage mapping to identify significant QTLs associated with grain quality traits in DH and F<sub>3</sub> tropical maize populations grown in Kenya. Numerous studies have extensively explored the genetic underpinnings of maize grain quality traits, resulting in the detection of a multitude of QTL (Li et al., 2009; Guo et al., 2013; Galić et al., 2017; Karn et al., 2017; Badu-Apraku et al., 2020; Ndlovu et al., 2022; Zhang et al., 2023).

In accordance with our results, the presence of QTLs related to oil content across all chromosomes except for chromosomes 2, 7, and 10, indicates a widespread genetic regulation of grain oil content in maize. Similarly, QTLs governing grain protein content were exclusively identified on chromosomes 1, 2, 3, 5, and 6, emphasizing the specific chromosomal regions contributing to protein variations. Furthermore, the near-ubiquitous distribution of QTLs associated with grain starch content across most chromosomes, except for chromosomes 9 and 10, highlights the genetic complexity underlying grain starch content in maize. These findings diverge slightly from those obtained by Zhang et al. (2023), who utilized different genotypes from the ones in our study and detected a combined total of 16 QTLs for grain oil content distributed across all maize chromosomes. Yang et al. (2012) also identified 58 QTLs for kernel oil content in all chromosomes. Interestingly, Zhang et al. (2023) recognized chromosome 9 as housing the largest effect QTLs for oil content. Yang et al. (2012) and Ndlovu et al. (2022), on the other hand, reported QTLs on chromosomes 1 and 2 as having the largest effects on grain oil content, respectively. Another study by Fang et al. (2021) identified five major effect QTLs associated with oil content located on chromosomes 6 and 9. For grain protein content, Lu et al. (2022) found associated QTLs in all chromosomes except chromosomes 6, 8 and 10. In the study by Ndlovu et al. (2022), the only major effect QTL associated with grain protein content was identified on chromosome 3. For starch content, the major effect QTLs on chromosomes 1, 3, 4, 5, 7, 8 and 9 were located in regions which were also reported in earlier studies (Goldman et al., 1993; Yang et al., 2013; Lin et al., 2019; Zhang et al., 2022). The absence of QTLs on certain chromosomes in our study, despite earlier reports, underscores the intricate interplay of genetic variations and environmental factors in QTL identification, likely influenced by the distinct genotypes and growing conditions used. This also highlights the importance of broadening the scope of grain quality trait breeding research to encompass diverse maize populations and environmental settings so as to comprehensively uncover QTLs associated with grain quality traits.

DH pop1 and F<sub>3</sub> pop3 share common parents. However, the average performance was relatively higher for most of traits in DH population with mean of 5.62%, 8.21%, 70.9% and 65.36 g observed for grain oil content, protein content, starch content and grain weight, respectively, compared to 5.17%, 7.33%, 70.9% and 54.3 g in F<sub>3</sub> pop3 (Figure 1). The number of QTL detected was also varied between these populations with more QTLs (4) detected for oil content in DH population whereas more QTL for starch content were detected in F<sub>3</sub> pop3 (Table 3). Interestingly, there is no overlapping QTLs were detected in this study. All QTL detected in F<sub>3</sub> pop3 are associated with minor effect whereas, with DH population, we able to identify one major effect QTL for grain protein content (*qGPC9-144* with PVE of 11.88%) and two major effect QTL for oil content (*qOC3-146* with PVE of 18.06%; *qOC6-102* with PVE of 26.85%). Overall, DH population has few advantages over F<sub>3</sub> population both in terms of mean performance and finding QTLs, however, results from different genetic populations are required for appropriate conclusion.

The QTLs identified in the present study were found to overlap with those previously reported in diverse maize populations (Ertiro et al., 2022; Ndlovu et al., 2022), indicating a degree of genetic consistency across different studies. A significant finding of our

analysis is the identification of distinct genomic regions on chromosome 6 associated with both oil and starch content (Table 3). The detection of QTLs in the genomic region spanning 89.81–112.9 Mb (chr 6) for these traits holds promise for further improvement of these traits. Furthermore, another distinct region from 188.45 to 233.61 Mb on chromosome 2 was found to harbour QTLs linked to grain protein and starch content. The identified regions on chromosomes 2 and 6 hold significant potential for enhancing grain quality in tropical maize through targeted breeding. Elucidating the underlying genes and mechanisms within these regions could provide invaluable insights for developing effective improvement strategies. In previous investigations, research studies by Wassom et al. (2008) and Fang et al. (2021) highlighted the crucial involvement of chromosome 6 in shaping grain quality traits, further substantiating its importance in genetic studies related to maize grain quality. In a similar context, earlier studies (Ndlovu et al., 2022; Zhang et al., 2022) have identified chromosome 2 as a pivotal genomic region associated with grain quality traits.

In maize, starch content is regulated by many genes (Zhong et al., 2020). Starch synthesis in kernels involves a series of starch metabolic enzymes like sucrose synthase (*SUS*), starch synthases (*SSs*) and starch branching and debranching enzymes (Nelson and Pan, 1995; Zhang et al., 2022). Finding candidate genes within the identified QTL regions improves the consistency of identified QTLs to be used in breeding to improve the linked trait. Two QTLs for starch content on F<sub>3</sub> pop2 (*qSC1-20* and *qSC1-21*) were overlapped with two QTLs for grain texture (*qGT1-43* and *qGT1-55*) and one QTL each for oil content (*qOC1-64*) and protein content (*qGPC1-105*) (Table 3). By using four DH populations, Zhang et al. (2022) also reported QTL for starch content in the same region, which is known to harbor sucrose synthase gene. *Sus2* is one of the three *SUS* encoding genes in maize located at 57.45 Mbp on chromosome 1, which has a unique role in cytoplasmic sucrose metabolism (Deng et al., 2020). Two QTLs from F<sub>3</sub> pop2 for starch content (*qSC4-10* and *qSC4-20*) are co-located with grain texture QTL (*qGT4-10*). Several studies also reported colocalized QTL in the same region (Wang et al., 2010; Guo et al., 2013; Yang et al., 2013; Dong et al., 2015; Hu et al., 2021; Zhang et al., 2022). This region also harbors three candidate genes involved in starch synthesis like vacuolar invertase (*IVR2*) at 69.45 Mbp, Starch branching enzyme (*SBE1*) at 65.19 Mbp and *SS2c* at 34.23 Mbp (Zhang et al., 2022). *IVR2* (Kim et al., 2000) in maize has a key role in carbon metabolism in both source and sink tissues that irreversibly hydrolyze sucrose to fructose and glucose and regulates sugar accumulation in sink organs (Juarez-Colunga et al., 2018). *SBE1* is related to amylose content and starch molecular structure (Zhong et al., 2021). Another candidate gene *SS2c* encodes soluble starch synthases which are responsible for synthesis of amylopectin (Yan et al., 2009). These results suggested that QTLs detected in this study are overlapped between traits and with earlier studies and related to series of candidate genes encoding key enzymes relevant to grain quality traits. Therefore, these QTLs are relevant and useful to be used in breeding through marker-assisted breeding to improve the grain quality in maize.

#### 4.3 Genomic selection in maize breeding for improved grain quality can serve as a valuable complement to conventional phenotypic selection methods

Genomic selection (GS) is rapidly gaining prominence in maize breeding programs, enabling the precise prediction of breeding values for individual maize lines (Crossa et al., 2017; Atanda et al., 2021; Singh and Kaundal, 2023). This approach has been extensively employed across diverse maize genotypes in numerous studies (Badji et al., 2020; Beyene et al., 2021; Gowda et al., 2021; Ma and Cao, 2021) to investigate various grain-related traits. Our genomic selection analysis revealed a range of prediction accuracies across the maize populations (Figure 3), indicating the promise of genomic information in predicting and selecting grain quality traits in maize breeding programs. Recorded moderate to high prediction accuracies offer promise in maize breeding for targeted traits (Gowda et al., 2021). Interestingly, our analysis revealed higher prediction accuracies for kernel weight, grain texture, and grain protein content, highlighting their suitability for GS. Additionally, we observed that overall trait prediction accuracy in DH pop 1 was superior as compared to the other three F<sub>3</sub> populations, indicating the potential for leveraging this DH population for more effective trait prediction and selection. A similar trend was observed by Ndlovu et al. (2022), who reported the highest prediction accuracies in DH populations under low N stress conditions. The differences between maize genotypes in overall trait accuracies underscore the importance of considering population-specific factors such as genetic diversity, population structure and trait variability within population or environmental interactions, which may influence the effectiveness of GS for grain quality traits.

## 5 Conclusion

Here, we utilized DH and F<sub>3</sub> tropical maize populations to detect significant QTLs associated with grain quality traits in maize lines grown in Kenya. The results of this investigation offer valuable genetic resources for molecular breeding and enhance our understanding of the genetic framework governing grain quality in tropical maize. The detection of multiple QTLs distributed across the 10 maize chromosomes suggests that the regulation of grain quality traits involves a combination of both major-effect and minor-effect QTLs. Notably, there was an observation of QTL overlap for protein, oil and starch contents, suggesting potential genetic links or pleiotropic effects. Moving forward, it is essential to prioritize the validation of these identified QTLs, as they hold the key to enhancing the efficiency of maize breeding programs aimed at improving grain quality within the SSA region. The results of this study also illustrated that integrating GS into tropical maize breeding programs focused on improving grain quality can serve as a valuable complement to conventional phenotypic selection methods.

## Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material. The marker data is accessible through <https://hdl.handle.net/11529/10548940> and <https://zenodo.org/records/10021760>.

## Author contributions

NN: Data curation, Formal Analysis, Investigation, Methodology, Resources, Software, Writing—original draft. RK: Data curation, Formal Analysis, Validation, Visualization, Writing—original draft. YB: Conceptualization, Funding acquisition, Project administration, Resources, Writing—review and editing. BD: Conceptualization, Project administration, Resources, Writing—review and editing. VO: Data curation, Formal Analysis, Writing—review and editing. DM: Conceptualization, Data curation, Writing—review and editing. CS: Conceptualization, Project administration, Resources, Supervision, Writing—review and editing. PM: Methodology, Visualization, Writing—review and editing. BP: Funding acquisition, Project administration, Resources, Supervision, Writing—review and editing. MG: Conceptualization, Formal Analysis, Funding acquisition, Methodology, Project administration, Software, Supervision, Validation, Writing—review and editing.

## Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. The research was supported by the Bill and Melinda Gates Foundation (B&MGF), Foundation for Food and Agriculture Research (FFAR) and the United States Agency for International Development (USAID) through the Stress Tolerant Maize for Africa (STMA, B&MGF Grant # OPP1134248) Project, IMAS Project, AGG (Accelerating Genetic Gains in Maize and Wheat for Improved Livelihoods;

## References

- Alvarado, G., Rodríguez, F. M., Pacheco, A., Burgueño, J., Crossa, J., Vargas, M., et al. (2020). META-R: a software to analyze data from multi-environment plant breeding trials. *Crop J.* 8, 745–756. doi:10.1016/j.cj.2020.03.010
- Álvarez-Iglesias, L., Malvar, R. A., Garzón, R., Rosell, C. M., and Revilla, P. (2021). Nutritional value of whole maize kernels from diverse endosperm types and effects on rheological quality. *Agronomy* 11, 2509. doi:10.3390/agronomy11122509
- Atanda, S. A., Olsen, M., Burgueño, J., Crossa, J., Dzidzienyo, D., Beyene, Y., et al. (2021). Maximizing efficiency of genomic selection in CIMMYT's tropical maize breeding program. *Theor. Appl. Genet.* 134, 279–294. doi:10.1007/s00122-020-03696-9
- Badji, A., Machida, L., Kwemoui, D. B., Kumi, F., Okii, D., Mwila, N., et al. (2020). Factors influencing genomic prediction accuracies of tropical maize resistance to fall armyworm and weevils. *Plants* 10, 29. doi:10.3390/plants10010029
- Badu-Apraku, B., Adewale, S., Angelot Paterne, A., Gedil, M., Toyinbo, J., and Asiedu, R. (2020). Identification of QTLs for grain yield and other traits in tropical maize under Striga infestation. *PLoS ONE* 15, e0239205. doi:10.1371/journal.pone.0239205
- Beyene, Y., Gowda, M., Pérez-Rodríguez, P., Olsen, M., Robbins, K. R., Burgueño, J., et al. (2021). Application of genomic selection at the early stage of breeding pipeline in tropical maize. *Front. Plant Sci.* 12, 685488. doi:10.3389/fpls.2021.685488
- Biswas, D. K., and Ma, B.-L. (2016). Effect of nitrogen rate and fertilizer nitrogen source on physiology, yield, grain quality, and nitrogen use efficiency in corn. *Can. J. Plant Sci.* 96, 392–403. doi:10.1139/cjps-2015-0186
- Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., and Buckler, E. S. (2007). TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics* 23, 2633–2635. doi:10.1093/bioinformatics/btm308
- Cimmyt, (2005). *Laboratory protocols: CIMMYT applied molecular genetics laboratory*. Third Edition. Mexico, D.F: CIMMYT. ISBN: 968-6923-30-6.
- Crossa, J., Pérez-Rodríguez, P., Cuevas, J., Montesinos-López, O., Jarquín, D., De Los Campos, G., et al. (2017). Genomic selection in plant breeding: methods, models, and perspectives. *Trends plant Sci.* 22, 961–975. doi:10.1016/j.tplants.2017.08.011
- Das, B., Atlin, G. N., Olsen, M., Burgueño, J., Tarekge, A., Babu, R., et al. (2019). Identification of donors for low-nitrogen stress with maize lethal necrosis (MLN) tolerance for maize breeding in sub-Saharan Africa. *Euphytica* 215, 80–15. doi:10.1007/s10681-019-2406-5
- De Groote, H., and Kimenju, S. C. (2012). Consumer preferences for maize products in urban Kenya. *Food Nutr. Bull.* 33, 99–110. doi:10.1177/156482651203300203
- Dei, H. K. (2017). "Assessment of maize (*Zea mays*) as feed resource for poultry," in *Poultry science*, 1–32. doi:10.5772/65363
- Deng, Y. T., Wang, J. C., Zhang, Z. Y., and Wu, Y. R. (2020). Transactivation of *Sus1* and *Sus2* by *Opaque2* is an essential supplement to sucrose synthase-mediated endosperm filling in maize *Sus1* and *Sus2* by *Opaque2* is an essential supplement to sucrose synthase-mediated endosperm filling in maize. *Plant Biotechnol. J.* 18, 1897–1907. doi:10.1111/pbi.13349

B&MGF Investment ID INV-003439) Project and the CGIAR Research Program on Maize (MAIZE). NN, PM, and CS acknowledge funding from the European Commission DESIRA funded LEG4DEV Project [FOOD/2020/418-901] and the Science Foundation Ireland Principal Investigator Grant [13/IA/1820].

## Acknowledgments

The authors thank the CIMMYT field technicians for phenotypic evaluations and Laboratory technicians for sample preparation for genotyping. We are also grateful to LGC genomic labs in the UK for genotyping the populations and providing the marker information. More gratitude to CIMMYT scientists for their excellent and noble guide throughout the processes of this research.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2024.1353289/full#supplementary-material>



- Dong, Y. B., Zhang, Z. W., Shi, Q. L., Wang, Q. L., Zhou, Q., and Li, Y. L. (2015). QTL identification and meta-analysis for kernel composition traits across three generations in popcorn. *Euphytica* 204, 649–660. doi:10.1007/s10681-015-1360-0
- Ekpa, O., Palacios-Rojas, N., Kruseman, G., Fogliano, V., and Linnemann, A. R. (2018). Sub-Saharan African maize-based foods: technological perspectives to increase the food and nutrition security impacts of maize breeding programmes. *Glob. Food Secur.* 17, 48–56. doi:10.1016/j.gfs.2018.03.007
- Erenstein, O., Chamberlin, J., and Sonder, K. (2021). Estimating the global number and distribution of maize and wheat farms. *Glob. Food Secur.* 30, 100558. doi:10.1016/j.gfs.2021.100558
- Erenstein, O., Jaleta, M., Sonder, K., Mottaleb, K., and Prasanna, B. (2022). Global maize production, consumption and trade: trends and R&D implications. *Food Secur.* 14, 1295–1319. doi:10.1007/s12571-022-01288-7
- Ertiro, B. T., Das, B., Kosgei, T., Tesfaye, A. T., Labuschagne, M. T., Worku, M., et al. (2022). Relationship between grain yield and quality traits under optimum and low-nitrogen stress environments in tropical maize. *Agronomy* 12, 438. doi:10.3390/agronomy12020438
- Ertiro, B. T., Labuschagne, M., Olsen, M., Das, B., Prasanna, B. M., and Gowda, M. (2020). Genetic dissection of nitrogen use efficiency in tropical maize through genome-wide association and genomic prediction. *Front. plant Sci.* 11, 474. doi:10.3389/fpls.2020.00474
- Fang, H., Fu, X., Ge, H., Zhang, A., Shan, T., Wang, Y., et al. (2021). Genetic basis of maize kernel oil-related traits revealed by high-density SNP markers in a recombinant inbred line population. *BMC Plant Biol.* 21, 344–412. doi:10.1186/s12870-021-03089-0
- FAOstat (2022). *Food and agriculture organization of the united nations*. Rome, Italy: FAOstat Statistical Database. Available at: <http://www.fao.org/faostat/en/home> (Accessed December 16, 2022).
- FAOSTAT (2023). “Crops and livestock products data from 2021 statistics,” in *Food and agriculture data* (Rome, Italy: FAO).
- Galić, V., Franić, M., Jambrović, A., Zdunić, Z., Brkić, A., and Šimić, D. (2017). QTL mapping for grain quality traits in testcrosses of a maize biparental population using genotyping-by-sequencing data. *Poljoprivreda* 23, 28–33. doi:10.18047/poljo.23.1.5
- Ganal, M. W., Durstewitz, G., Polley, A., Bérard, A., Buckler, E. S., Charcosset, A., et al. (2011). A large maize (*Zea mays* L.) SNP genotyping array: development and germplasm genotyping, and genetic mapping to compare with the B73 reference genome. *PLoS one* 6, e28334. doi:10.1371/journal.pone.0028334
- Gilmour, A. R., Gogel, B. J., Cullis, B. R., Welham, S., and Thompson, R. (2002). *ASReml user guide release 1.0*.
- Goering, R., Larsen, S., Tan, J., Whelan, J., and Makarevitch, I. (2021). QTL mapping of seedling tolerance to exposure to low temperature in the maize IBM RIL population. *PLoS one* 16, e0254437. doi:10.1371/journal.pone.0254437
- Goldman, I., Rocheford, T., and Dudley, J. (1993). Quantitative trait loci influencing protein and starch concentration in the Illinois long term selection maize strains. *Theor. Appl. Genet.* 87, 217–224. doi:10.1007/BF00223767
- Goredema-Matongera, N., Ndhlela, T., Magorokosho, C., Kamutando, C. N., van Biljon, A., and Labuschagne, M. (2021). Multinutrient biofortification of maize (*Zea mays* L.) in Africa: current status, opportunities and limitations. *Nutrients* 13, 1039. doi:10.3390/nu13031039
- Gowda, M., Beyene, Y., Makumbi, D., Semagn, K., Olsen, M. S., Bright, J. M., et al. (2018). Discovery and validation of genomic regions associated with resistance to maize lethal necrosis in four biparental populations. *Mol. Breed.* 38, 66–16. doi:10.1007/s11032-018-0829-7
- Gowda, M., Makumbi, D., Das, B., Nyaga, C., Kosgei, T., Crossa, J., et al. (2021). Genetic dissection of *Striga hermonthica* (Del.) Benth. resistance via genome-wide association and genomic prediction in tropical maize germplasm. *Theor. Appl. Genet.* 134, 941–958. doi:10.1007/s00122-020-03744-4
- Guo, Y., Yang, X., Chander, S., Yan, J., Zhang, J., Song, T., et al. (2013). Identification of unconditional and conditional QTL for oil, protein and starch content in maize. *Crop J.* 1, 34–42. doi:10.1016/j.cj.2013.07.010
- Hammad, H. M., Abbas, F., Ahmad, A., Farhad, W., Wilkerson, C. J., and Hoogenboom, G. (2023). Water and nitrogen management influence on oil and protein concentration in maize. *Agron. J.* 115, 557–568. doi:10.1002/agj.21275
- Hirel, B., Le Gouis, J., Ney, B., and Gallais, A. (2007). The challenge of improving nitrogen use efficiency in crop plants: towards a more central role for genetic variability and quantitative genetics within integrated approaches. *J. Exp. Bot.* 58, 2369–2387. doi:10.1093/jxb/erm097
- Hu, S. T., Wang, M., Zhang, X., Chen, W. K., Song, X. R., Fu, X. Y., et al. (2021). Genetic basis of kernel starch content decoded in a maize multi-parent population. *Plant Biotechnol. J.* 19, 2192–2205. doi:10.1111/pbi.13645
- Karn, A., Gillman, J. D., and Flint-Garcia, S. A. (2017). Genetic analysis of teosinte alleles for kernel composition traits in maize. *G3 Genes, Genomes, Genet.* 7, 1157–1164. doi:10.1534/g3.117.039529
- Katsenios, N., Sparangis, P., Chanioti, S., Giannoglou, M., Leonidakis, D., Christopoulos, M. V., et al. (2021). Genotype × environment interaction of yield and grain quality traits of maize hybrids in Greece. *Agronomy* 11, 357. doi:10.3390/agronomy11020357
- Kim, J. Y., Mahé, A., Brangeon, J., and Prioul, J. L. (2000). A maize vacuolar invertase, IVR2, is induced by water stress. Organ/tissue specificity and diurnal modulation of expression. *Plant Physiol.* 124, 71–84. doi:10.1104/pp.124.1.71
- Kimutai, C., Ndlovu, N., Chaikam, V., Ertiro, B. T., Das, B., Beyene, Y., et al. (2023). Discovery of genomic regions associated with grain yield and agronomic traits in Biparental populations of maize (*Zea mays* L.) under optimum and low nitrogen conditions. *Front. Genet.* 14, 1266402. doi:10.3389/fgene.2023.1266402
- Kosambi, D. D. (1944). The estimation of map distance from recombination values. *Ann. Eugen.* 12, 172–175. doi:10.1111/j.1469-1809.1943.tb02321.x
- Langyan, S., Bhardwaj, R., Kumari, J., Jacob, S. R., Bisht, I. S., Pandravada, S. R., et al. (2022). Nutritional diversity in native germplasm of maize collected from three different fragile ecosystems of India. *Front. Nutr.* 9, 812599. doi:10.3389/fnut.2022.812599
- Li, Y., Wang, Y., Wei, M., Li, X., and Fu, J. (2009). QTL identification of grain protein concentration and its genetic correlation with starch concentration and grain weight using two populations in maize (*Zea mays* L.). *J. Genet.* 88, 61–67. doi:10.1007/s12041-009-0008-z
- Lin, F., Zhou, L., He, B., Zhang, X., Dai, H., Qian, Y., et al. (2019). QTL mapping for maize starch content and candidate gene prediction combined with co-expression network analysis. *Theor. Appl. Genet.* 132, 1931–1941. doi:10.1007/s00122-019-03326-z
- Liu, Z., Xie, H., Tian, G., Chen, S., Wang, C., Hu, Y., et al. (2008). QTL mapping of nutrient components in maize kernels under low nitrogen conditions. *Plant Breed.* 127, 279–285. doi:10.1111/j.1439-0523.2007.01465.x
- Lu, X., Zhou, Z., Wang, Y., Wang, R., Hao, Z., Li, M., et al. (2022). Genetic basis of maize kernel protein content revealed by high-density bin mapping using recombinant inbred lines. *Front. Plant Sci.* 13, 1045854. doi:10.3389/fpls.2022.1045854
- Ma, J., and Cao, Y. (2021). Genetic dissection of grain yield of maize and yield-related traits through association mapping and genomic prediction. *Front. Plant Sci.* 12, 690059. doi:10.3389/fpls.2021.690059
- Ndlovu, N., Spillane, C., McKeown, P. C., Cairns, J. E., Das, B., and Gowda, M. (2022). Genome-wide association studies of grain yield and quality traits under optimum and low-nitrogen stress in tropical maize (*Zea mays* L.). *TAG. Theor. Appl. Genet. Theor. Und Angewandte Genet.* 135, 4351–4370. doi:10.1007/s00122-022-04224-7
- Nelson, O., and Pan, D. (1995). Starch synthesis in maize endosperms. *Annu. Rev. Plant Biol.* 46, 475–496. doi:10.1146/annurev.pp.46.060195.002355
- Prasanna, B. M., Cairns, J. E., Zaidi, P., Beyene, Y., Makumbi, D., Gowda, M., et al. (2021). Beat the stress: breeding for climate resilience in maize for the tropical rainfed environments. *Theor. Appl. Genet.* 134, 1729–1752. doi:10.1007/s00122-021-03773-7
- Quarshie, P. T., Abdulai, A.-R., and Fraser, E. D. (2021). Africa’s “seed” revolution and value chain constraints to early generation seeds commercialization and adoption in Ghana. *Front. Sustain. Food Syst.* 5, 665297. doi:10.3389/fsufs.2021.665297
- Ranum, P., Peña-Rosas, J. P., and Garcia-Casal, M. N. (2014). Global maize production, utilization, and consumption. *Ann. N. Y. Acad. Sci.* 1312, 105–112. doi:10.1111/nyas.12396
- Ray, K., Banerjee, H., Dutta, S., Hazra, A. K., and Majumdar, K. (2019). Macronutrients influence yield and oil quality of hybrid maize (*Zea mays* L.). *PLoS one* 14, e0216939. doi:10.1371/journal.pone.0216939
- Ren, J., Wu, P., Huestis, G. M., Zhang, A., Qu, J., Liu, Y., et al. (2022). Identification and fine mapping of a major QTL (qRts8-1) conferring resistance to maize tar spot complex and validation of production markers in breeding lines. *Theor. Appl. Genet.* 135, 1551–1563. doi:10.1007/s00122-022-04053-8
- Ribaut, J.-M., Jiang, C., Gonzalez-de-Leon, D., Edmeades, G., and Hoisington, D. (1997). Identification of quantitative trait loci under drought conditions in tropical maize. 2. yield components and marker-assisted selection strategies: 2. yield components and marker-assisted selection strategies. *Theor. Appl. Genet.* 94, 887–896. doi:10.1007/s001220050492
- Singh, V., and Kaundal, A. (2023). “Genomic selection in maize breeding,” in *Maize improvement: current advances in yield, quality, and stress tolerance under changing climatic scenarios* (Springer), 99–108.
- Sitonik, C. a., Suresh, L., Beyene, Y., Olsen, M. S., Makumbi, D., Oliver, K., et al. (2019). Genetic architecture of maize chlorotic mottle virus and maize lethal necrosis through GWAS, linkage analysis and genomic prediction in tropical maize germplasm. *Theor. Appl. Genet.* 132, 2381–2399. doi:10.1007/s00122-019-03360-x
- Wang, L., Yu, X., Gao, J., Ma, D., Guo, H., and Hu, S. (2023). Patterns of influence of meteorological elements on maize grain weight and nutritional quality. *Agronomy* 13, 424. doi:10.3390/agronomy13020424
- Wang, Y. Z., Li, J. Z., Li, Y. L., Wei, M. G., and Fu, J. F. (2010). QTL detection for grain oil and starch content and their associations in two connected F2:3 populations in high-oil maize. *Euphytica* 174, 239–252. doi:10.1007/s10681-010-0123-1
- Wassom, J. J., Wong, J. C., Martinez, E., King, J. J., DeBaene, J., Hotchkiss, J. R., et al. (2008). QTL associated with maize kernel oil, protein, and starch concentrations; kernel mass; and grain yield in Illinois high oil × B73 backcross-derived lines. *Crop Sci.* 48, 243–252. doi:10.2135/cropsci2007.04.0205

- Worku, M., Bänziger, M., Friesen, D., Diallo, A. O., and Horst, W. J. (2012). Nitrogen efficiency as related to dry matter partitioning and root system size in tropical mid-altitude maize hybrids under different levels of nitrogen stress. *Field Crops Res.* 130, 57–67. doi:10.1016/j.fcr.2012.02.015
- Yan, H. B., Pan, X. X., Jiang, H. W., and Wu, G. J. (2009). Comparison of the starch synthesis genes between maize and rice: copies, chromosome location and expression divergence. *Theor. Appl. Genet.* 119, 815–825. doi:10.1007/s00122-009-1091-5
- Yang, G., Dong, Y., Li, Y., Wang, Q., Shi, Q., and Zhou, Q. (2013). Verification of QTL for grain starch content and its genetic correlation with oil content using two connected RIL populations in high-oil maize. *PLoS One* 8, e53770. doi:10.1371/journal.pone.0053770
- Yang, J., Liu, Z., Chen, Q., Qu, Y., Tang, J., Lübberstedt, T., et al. (2020). Mapping of QTL for grain yield components based on a DH population in maize. *Sci. Rep.* 10, 7086. doi:10.1038/s41598-020-63960-2
- Yang, X., Ma, H., Zhang, P., Yan, J., Guo, Y., Song, T., et al. (2012). Characterization of QTL for oil content in maize kernel. *Theor. Appl. Genet.* 125, 1169–1179. doi:10.1007/s00122-012-1903-x
- Zhang, L., Liang, Z.-y., He, X.-m., Meng, Q.-f., Hu, Y., Schmidhalter, U., et al. (2020). Improving grain yield and protein concentration of maize (*Zea mays* L.) simultaneously by appropriate hybrid selection and nitrogen management. *Field Crops Res.* 249, 107754. doi:10.1016/j.fcr.2020.107754
- Zhang, X., Wang, M., Guan, H., Wen, H., Zhang, C., Dai, C., et al. (2023). Genetic dissection of QTLs for oil content in four maize DH populations. *Front. Plant Sci.* 14, 1174985. doi:10.3389/fpls.2023.1174985
- Zhang, X., Wang, M., Zhang, C., Dai, C., Guan, H., and Zhang, R. (2022). Genetic dissection of QTLs for starch content in four maize DH populations. *Front. Plant Sci.* 13, 950664. doi:10.3389/fpls.2022.950664
- Zheng, Y., Yuan, F., Huang, Y., Zhao, Y., Jia, X., Zhu, L., et al. (2021). Genome-wide association studies of grain quality traits in maize. *Sci. Rep.* 11, 9797–9812. doi:10.1038/s41598-021-89276-3
- Zhong, Y. Y., Liu, L. S., Qu, J. Z., Li, S. L., Blennow, A., Seytahmetovna, S. A., et al. (2020). The relationship between the expression pattern of starch biosynthesis enzymes and molecular structure of high amylose maize starch. *Carbohydr. Polym.* 247, 116681. doi:10.1016/j.carbpol.2020.116681
- Zhong, Y. Y., Qu, J. Z., Blennow, A., Liu, X. X., and Guo, D. W. (2021). Expression pattern of starch biosynthesis genes in relation to the starch molecular structure in high-amylose maize. *J. Agric. Food Chem.* 69, 2805–2815. doi:10.1021/acs.jafc.0c07354